Deterrent Effects of Warnings on User's Behavior in Preventing Malicious Software Use

Mario Silic
University of St Gallen, Switzerland
mario.silics@unisg.ch

Andrea Back
University of St Gallen, Switzerland
andrea.back@unisg.ch

Abstract

Despite the fact that a number of technical counter-measures do exist to mitigate the risks related to malicious software, in reality users are the last line of defense against security incidents. In this technology-human interaction, warning messages can represent an important tool to help users when making a decision. Understanding the effects of computer warnings on the progression and duration of the malicious software use would bridge the existing knowledge gap. Supported by the restrictive deterrence model and psychological factors, we conducted a non-controlled field experiment in which we collected data from no previously recruited participants. We found that in the presence of the warning message, the progression of the software use will be decreased and the duration of both first and repeated software uses will be reduced. Finally, we offer important findings for further theorizing and interesting practitioner insights that could help to leverage the interaction between the human and the computer technology with an objective to reduce the risk.

1. Introduction

The recent security incidents of Target, Home Depot and Sony Pictures revealed how destructive malware (malicious software) can be to the organizational reputation and financial stability. An organization can receive an average of nearly 17,000 malware alerts in a typical week which represents a significant amount of time to respond to these alerts impacting organization's financial resources and IT personnel [1]. It is estimated that over 800 million people suffered from security incidents (e.g. stealing user's private information) in 2013 [2]. Although, a number of technical counter-measures do exist to mitigate the risks (e.g. personal anti-virus), in reality users represent the last gate in the decision making process. In this technology-user interaction, warning messages can represent an important tool to help them when taking a decision [3]. Warnings represent communications designed to prevent users from hurting themselves or others [4, 5]. Clearly, warning is not the best option as the decision task is on the user who has to make a choice. Consequently, users who are constantly exposed to security warnings [6] often ignore them due to habituation [3, 7, 8].

Past research has tried to better understand how users interact with warnings and why users ignore them [9-13]. Mostly the focus was on examining the SSL web browser warning messages and their effectiveness [e.g. 14, 15-17] where participants were directly recruited. This might present a bias as the population was not randomly chosen which could lead to users being more likely to click through warning dialog messages and less concerned about their own privacy [14]. Moreover, little research has examined the effects of warnings on the progression of the incident event related to the malware use.

Precisely, most of the studies simply examined if user heeds the warning message by measuring user's decision which, mostly, resulted in the binary outcome: continue or exit. In this context, it is unclear if the warning message has any effect on user's decision.

Interestingly, while number of studies have used deterrence theory to understand how fear of sanctions and punishments prevent deviance and crime, only few studies have studied the effect of punishment threats in reducing the frequency and severity of individual offending as suggests restrictive deterrence theory [18]. Restrictive deterrence suggests that an individual who commits an act of crime at least once will be mainly preoccupied with reduction in the frequency of the illegal act [18]. More precisely, an offender, knowing that the criminal act is committed, will seek to decrease the frequency of its offending hoping to avoid being caught. Restrictive deterrence concept is particularly useful in our context as it allows to understand the link between the presence of sanction threats (e.g. warning message communication) and the restriction of the scope of user's illegal activities (e.g. reducing the frequency of the malicious software use). Past studies have failed

URI: http://hdl.handle.net/10125/41762
ISBN: 978-0-9981331-0-2

to clearly establish this link between the warning message communication and the frequency, duration and progression of an event [19].

By addressing this challenge, we aim to bridge this gap by examining how warnings can lead to a higher effectiveness of sanction threats in presence of malware when it comes to the progression, reduction in frequency and decrease in duration of the malware use.

Supported by the deterrence theory, and in particular the restrictive deterrence model as suggested by [18, 20], we aim to investigate three research questions: 1) is the warning message leading to an immediate incident termination?; 2) is the warning message impacting the frequency of repeated malware use? and 3) does the warning message affect the duration of the hazard?

2. Theoretical foundation

2.1. Restrictive deterrence

Deterrence theory, a prominent theory from the Criminology field, suggests that individuals that intend to commit a crime or antisocial act can be dissuaded by the implementation of sanctions and disincentives that are relevant to these acts [21]. Theory posits that there is a high chance of being caught and punished severely. This general deterrence theory has been extended, in the last years, by contemporary theoreticians who proposed the 'restrictive deterrence' model which represents the process whereby offenders limit the frequency and severity of individual offending [18, 20]. For Gibbs [18] restrictive deterrence can be defined as "the curtailment of a certain type of criminal activity by an individual during some period because in whole or in part the curtailment is perceived by the individual as reducing the risk that someone will be punished as a response to the activity" (1975: 33). Surprisingly, little research has examined the restrictive deterrence aspects and its relationship with deterring user from committing risky or bad actions [22]. Paradoxically, in the malware context, the offender is the user itself who is confronted by the deterrent warning message informing the user about possible sanctions that he or she may incur if the action is continued. Therefore, it is expected, according to the restrictive deterrence, that the user will reduce the frequency of its acts as user will be sanctioned at some point in time.

Interestingly, most of the past studies that examined the restrictive deterrent concept were using the qualitative research method [e.g. 20, 23-25] investigating a relatively small samples [e.g. 26]. One important reason for this lack of the quantitative studies could be the access to data as not only that it is difficult to build a study that would deal with the malware context but also, how to avoid bias by not recruiting participants directly.

2.2. Warnings, human interaction and restrictive deterrence

Restrictive deterrence theory suggests that in the malware context, users should reduce the frequency and duration of their acts. In order to communicate the sanction threat, the warning message is commonly used as the communication medium through which, hazard is explained. However, relying on the individual, to take the ultimate decision whether to comply or not, is not the best option.

Indeed, hazard and control hierarchy model [4] suggests that warning should only be the third option presented to the user. The model argues that the first step is to try to eliminate or remove hazard as much as possible. The second step should be the avoidance of the interaction between the user and the hazard and last option should be to present the warning to the user who will eventually take the decision.

To better understand this interaction between the humans and the technology, several models and framework have been suggested. The human-in-the-loop (HITL) framework was proposed as a general model suggesting a systematic approach to identify potential causes for human failure [27]. This model is based on the communication-human information processing (CHIP) model that describes the processing steps that are undertaken by the user when confronted by the warning message [28]. The mental model distinguished between novice and advanced users that make sense of warnings in different ways, coming to different conclusions and consequently, respond and act differently [29]. Overall, these models try to explain the interaction between humans and technology that is sequenced and split in different stages with the end goal to change the user's behavior.

User's ignorance of warnings is explained by the fact that users have difficulties distinguishing the real threats from the false ones [17]. However, the effectiveness of warnings can be increased by a warning text that includes a clear and non-technical description of potential negative outcome [16]. Also, positioning of the dialog warning message, the amount of text, the content length, manipulation of the warning content and the amount of technical details are some of the cues used to draw the user's attention [15, 30, 31]. One issue with these studies is

that they used a 'trial and error' approach in building the warning content [16] instead of designing the content based on some theoretical foundations.

When it comes to the restrictive deterrence premises, where user behavior is expected to be influenced by the warning message communication, we are still missing a better understanding of this relationship and how the interaction between the user and the warning will influence the decision making process. Indeed, while majority of past studies focused on understanding the effectiveness of warnings in preventing and deterring the occurrence of the hazard related to the malware propagation and use, little has been done to investigate the impact of warnings on the *progression* and *duration* of the hazard. In agreement with [18, 20, 22], we argue that while the malware hazard may not be fully stopped due to the warning, understanding the occurrence and progression of the hazard is of theoretical importance. Indeed, we expect to see users behaviors impacted by the warning content built around psychological cues informing the user about a clear description of potential negative outcome that the user will incur in case of the non-compliance.

Hence, based on the restrictive deterrence model and supported by the psychological factors, we explore this interaction between the human and the technology and how the warning message, in relation to the progression and duration of hazard, influences the human decision.

3. Research hypothesis

Recent reports on malware progression suggest that "deceptive downloads" are currently ranked as major threat as "deceptive downloads pretend to be installers for legitimate software but actually steal or destroy the user's sensitive and/or valuable data" [32]. The recommendation provided to end users is to immediately stop using the potentially malicious software and quit it. For example, Apple suggests to quit Safari web browser or cancel the installation process if malware has been downloaded [33].

Past studies reported mixed and often inconclusive results that were dependent on how the warning message was designed. For instance, it was found that users react differently when warning is displayed by Firefox web browser compared to Chrome [14]. In our study, we are looking to understand the effectiveness of the displayed warning message in determining the progression of the hazard, and thereby to explore if the user's behavior will lead to the decreased software use. Precisely, the impact of the warning message will be of high importance to user's decision making process only if warning message is able to capture users' attention and convey information about the possible hazard [17].

According to Wogalter and Laughery [34], user's attention will be driven by (1) spatial and temporal factors such as novelty, size, illumination, and contrast, (2) signal words such as "DANGER", (3) signal icons such as an exclamation point, (4) color such as red which signals danger in many cultures, and (5) pictures such as a pictorial sign displaying smoking consequences. One study on web browser warnings, such as those that appear when users visit suspected phishing websites, showed that altering text and color led to a significant increase of user's attention [31]. Clearly, if user is communicated the risk that malware will, for instance, destroy its hard disk, the chances are much higher that user's attention and consequently, its decision will be affected. In this context, we would expect that higher impact-warning message (i.e. warning message that, for instance, communicates direct risk for user's data will lead to a decreased software use.

Therefore, we hypothesize that the progression of the software use, in presence of the warning message, will be positively influenced and decreased by the higher impact-warning message

In line with the restrictive deterrence model [18], we argue that during a repeated use of the software, users will pay more attention to the warning and will be less inclined to ignore it, which should lead to a decreased use and consequently, to abandonment. Specifically, it means that users may feel that at a certain point in time they will experience negative consequences from their act. Hence, we can expect that users will try to decrease or avoid their risky by reducing their frequency.

Therefore, we hypothesize that *the frequency of repeated software use will be decreased in the presence of the warning message*.

Moreover, according to Jacobs [20] the warning message threat will lead to the restriction of the scope of the user's behavior. This suggests that when user is presented with a warning message for the first time, and if the message is displayed during the repeated software use, duration and progression of the hazard should be reduced. When hackers try to access an unauthorized system a repeated warning may decrease the duration of the security incident related to their system use, especially if they believe their actions are monitored [22]. This "limiting exposure" factor is also highlighted in many counter-measure practitioner suggestions [e.g. 33] and we argue that in the malware context, users will seek to limit and shorten the hazard time.

Therefore, we hypothesize that *the warning message will reduce the duration of both first and repeated software uses.*

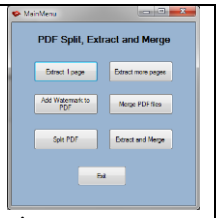
4. Research methodology

We designed a non-controlled field experiment where users can freely download software to manipulate PDF documents from different open source web repositories (e.g. sourceforge, github, etc.). Software was created by one of the authors using Microsoft Visual Basic programming language. Software (name PDF Split, Extract and Merge) is a fully functional application that allows users to manipulate PDF documents (e.g. split, merge, extract).

The reason for choosing to build PDF is because PDF software was found to be one of the most used rogue IT categories within organizations [35].

On Figure 1 application screenshots are presented.

a) start up screen



b) main menu

Figure 1. Software used for the field experiment

Once the user downloads the application and runs it for the first time the startup screen (Figure 1.a) will appear where user has to click on start button. After clicking on the button, user has to read, accept EULA license and provide his or her consent for participating in the research study. Once user agreed, one of the four warnings will be displayed to the user: 1) no-warning message (used as control group); 2) warning type 1 – low impact; 3) warning type 2 – medium impact and 4) warning type 3 – high impact. Each of the three warning messages (except the control message) expresses different consequences for the user if user continues to use the software that can be against general security policies , software use is illegal and monitored, or it is potentially dangerous and malicious .

We use the "no-warning" message as a control group to understand the impact of other warnings. The three warning messages design is based on the McAfee Security Center layout. McAfee Security Center is the graphical user interface for other McAffee products such as McAfee Antivirus. We kept the original layout but adapted the text for each of the three warning messages. The low impact warning message does not have any explicit warning design elements and is simply advising user not to continue using the application. The medium warning contains a more explicit warning message informing the user about the legal sanctions and the fact that software use is monitored and tracked. Finally, the high impact warning message communicates a clear risk for the user ("This application can be dantgerous! It can damage your hard disk and erase all your data!").

All three textual elements are based on past studies which used the same or similar textual content in various contexts [9, 17, 22, 30]. For instance, Maimon, Alper, Sobesto and Cukier [22] used similar content to inform hackers about risks they incur if they penetrate organizational systems.

Display of the warnings is randomized and controlled by the random function within the software.

In each of the cases, the user is presented with two options: exit or continue.

We record four measures:

- *decision* that is recorded as "0" if use clicks on 'Exit' button and "1" if user chooses "Continue"
- *duration* is recorded in milliseconds and represents the time passed between the click on 'Start' (Figure 1a) and the decision (exit or continue)
- *IP address*: user's internet address (IP address) is recorded and is used to have more insights on users' country origin
- *MAC address* that represents the unique identifier of each PC, assigned to each network device (e.g. network card). This measure is used to understand the frequency and repeated software use which can reveal repeated user's behavior

Progression of software use consists of the total duration (in milliseconds) representing the time user spent deciding whether to continue or exit. Progression is operationalized through the duration measure. Frequency of repeated software use is defined by the subsequent software uses where frequency can be 1 if, for example, user used software only once or it can be more than 1 suggesting that user continued to use software despite the warning message presence. Frequency is operationalized through the UserID measure (unique value assigned to each user and combination of MAC and IP addresses). Duration of first and repeated software uses corresponds to the duration (in milliseconds) where user's initial software use will be registered as first (UserID is stored in the database)

and repeated software use are all subsequent registered software uses (e.g. second use will have a certain duration time).

4.1. Participants

As the software was placed on the internet, any user was able to download, install and use the software freely. This means that we did not recruit any participants for the study, which increases the study's validity. By doing so, we were able to create and simulate a genuine environment.

Institutional Review Board (IRB) approval was given to collect data and human-subject protocols were followed. In addition, every participant had to provide his or her consent for being part of a research study. Once the application was started, a dialog box was opened informing the user about study's objectives (and informing them that no identifiable information would be collected) and asking them to confirm their participation. If users' would chose not to participate then we would not measure any of their activities (this was set programmatically). Hence, users were fully aware of the experiment.

Also, all participants had to accept end user license agreement (EULA) which, among other clauses, stipulated that "By downloading this software, you consent to send usage information to improve this product and future research".

4.2. Results

While we did not have any demographics collected, as we did not actively recruit participants, the only information that was available is the user's country. In total, 790 events were recorded (in 35 cases users chosen not to participate in the study – for them we did not collect any information but just counted their refusal to participate). We had users from 75 different countries that downloaded the software at least once. In table 1 the breakdown of the top 20 users' country downloads is presented.

Country	**% of downloads**	**Country**	**% of downloads**
United States	29%	Indonesia	2%
Germany	13%	Mexico	2%
India	13%	United Kingdom	2%
Italy	7%	Poland	2%
Spain	4%	Sweden	2%
France	3%	Brazil	2%
Netherlands	3%	China	2%
Turkey	3%	Russia	2%
Canada	2%	Australia	1%
Romania	2%	Singapore	1%

Table 1. Top 20 users' country downloads

In Table 2 a detailed overview of warnings displayed and the corresponding user actions can be found. Exit action was chosen in 36% of all cases, while 64% of users decided to continue with the software use. When it comes to the warning types, as expected, for the 'No warning' message only few users (10%) stopped using the application while the large majority (90%) continued. Other warning types (low and medium impact warning message) had similar results where majority of users continued their behavior and were not influenced by the risk suggested by the warning. However, the 'high warning' message seemed to have a different effect where 63% of users found the message to be rather persuasive and thus, decided to exit the software use compared to the 37% of users who continued.

Overview of all warnings types			
Warning type	**Exit action** (decision=0)	**Continue action** (decision=1)	**Total**
Low Warning	78 (34%)	151 (66%)	229
Medium Warning	95 (44%)	121 (56%)	216
High Warning	90 (63%)	53 (37%)	143
No Warning	19 (10%)	179 (90%)	198
Warning vs Control			
Warning	266 (45%)	326 (55%)	592
Control	19 (10%)	179 (90%)	198
Grand Total	284 (36%)	506 (64%)	790

Table 2. Overview of warnings display and users' actions

We conclude that hypothesis 1 is supported as progression of the software use, in presence of the warning message, will be positively influenced and decreased by the higher impact-warning message.

Next, we wanted to understand if frequency of repeated software use will be decreased in the presence of the warning message. We used the information from the entire set of incidents (N=790) and estimated whether the mean number of repeated

incidents is significantly different between the warning and the control group. Hence, we test for a significant difference between the proportions of immediate cessation on warning (when user is presented with one of the three warnings) event and no-warning (control group). We performed a t-test to compare the two proportions using decision as a dependent variable. The results from this test revealed an insignificant difference between these two groups ($Z= -6.6$; $p > .05$).

We conclude that hypothesis 2 is not supported as the frequency of repeated software use is not decreased in the presence of the warning message.

Further, to understand how warning message impact the duration and the progression of the hazard, we analyze the survival time of the software use. As we cannot simply compare the average hazard durations, due to the right skewed distribution of the survival time, we use event history analysis technique. One such survival technique is the Kaplan-Meier Survival estimator [36] which enables dealing with differing survival times (times-to-event), especially when not all the subjects continue in the study [37].

The survival rate is expressed as the survivor function (S): where t is a time period known as the survival time, time to failure or time to event (e.g. exit or continue action).

The results of the Kaplan-Meier estimate are presented on Figure 2. We can see that warning messages have high effect on the survival times between the control warning message (no warning message) and all other warning messages. It means that the proportion of the software use that survived is shorter on the treatment (warning) than on the control (no-warning).

By estimating the Cox proportional-hazard regression, we tested the significance of the effects of warning on the hazard duration. The Cox model allows investigation of the relationship between the survival of the event and independent measures [38].

The results, calculated using mplus software, are presented in Table 3.

Results confirm that the warning banner has a positive association with the hazard of the first observed event termination.

Hence, the hazard ratio estimate of the warning measure shows that the warning message is significantly (more than 1.6 times) increasing the rate of first observed event leading to much shorter event duration.

Consequently, hypothesis 3 is supported as hazard ration increases in presence of the warning message, indicating that the warning message will reduce the duration of both first and repeated software uses.

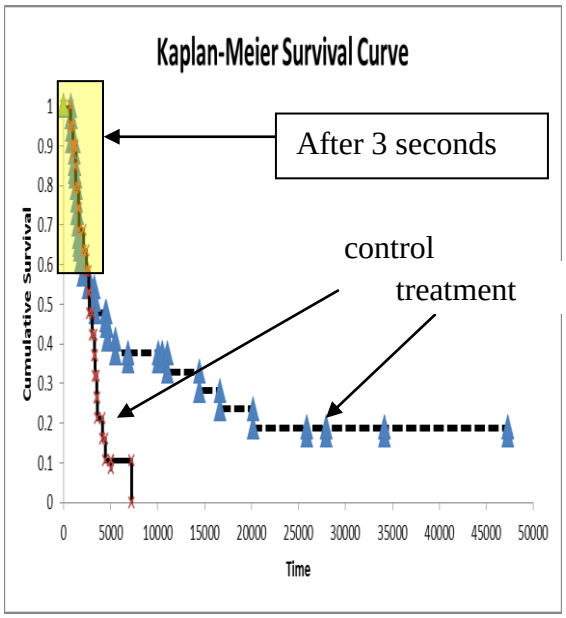


Figure 2. Kaplan-Meier function (in milliseconds)

	Coefficient (standard error)	**Hazard ratio**	**Log Likelihood**
First observed events (N=248)	0.75*	1.65	-120.23
All observed events (N=790)	0.181*	0.774	-1541.356
∗p < .05 (two-tailed);			

Table 3. Cox Proportional hazards survival regression results

5. Discussion

Our paper makes several contributions. First, we studied the interaction between the human and the technology in the context of malware use. Specifically, how users react in the presence of sanctions highlighted by the warning message. Indeed, computer users in their communication and interaction process with the technology will be more inclined to follow the procedure if there is a security concern behind [39], but their decision making process may be burdened by the overwhelming

amount of warnings remains [16]. This is very true in many contexts where users are frequently seeing too many warnings. For example, users clicked through Google Chrome's SSL warning 70.2% of the time but only 33.0% of Mozilla Firefox's SSL warnings [14]. Our study examined three hypotheses.

First, we hypothesized that the progression of the software use, in presence of the warning message, will be positively influenced and decreased by the higher impact-warning message. Unlike the past studies which often had mixed and inconclusive results, we clearly found support for this hypothesis where the progression of the software use is impacted by the warning message. Furthermore, as a higher degree of impact is communicated to the user, lower click-through is observed. This finding bridges the gap of the past studies, which mostly focused on one single warning type and in a particular environment (e.g. web browser). Interestingly, during the first three seconds, the proportion of survival events is very similar for both warning and no-warning context. This could be explained by the fact that users are simply habituated to see warnings and simply ignore them [3, 8] or it can be that, as suggested by the mental model approach [40], users can be more advanced in terms of their technological skills and thus, can better assess the risks than the novices users.

Other possible explanations are that users did not read the computer warning [7], did not understand the warning [41] or simply do not heed them [42]. In this content, our finding is in line with these explanations. However, we clearly show that the click through (CTR) decreases with the impact level of the warning message. From the low impact warning (66% CTR), medium (56% CTR) to high impact (37% CTR), there is a clear impact on the progression and duration of the software use. It would be very interesting to see what would happen if a highly effective (warning message communicating high risk) warning message would be communicated to users. Would the click through decrease to a very low and acceptable level so we could confirm that warning message leads to an immediate cessation of the software use?

Second, we hypothesized that the frequency of repeated software use will be decreased in the presence of the warning message. Contrary to the suggestion of the restrictive deterrence model [18], we did not find any support that the frequency of repeated uses is reduced in presence of the warning message. More precisely, average number of times user is using software is not different in presence of the warning or the control message. Hence, the frequency of repeated uses is not affected and user decision-making process remains consistent whether in presence of the warning or non-warning (i.e. control) message.

We explain this by the fact that if the user is at first presented with the high impact warning message and user ignores it, all the successive uses will also not be deterred as if user was not deterred by the high impact warning, why would he or she care about the lower impact warning content? However, if it was the opposite scenario where user saw the low warning message that was followed by medium or high impact warning message, it could be that user's frequency of repeated uses would be reduced. This is something that was found to be effective in the tobacco industry where users instead of being presented with the standard packaging message (e.g. "Smoking can cause a slow and painful death"), would be presented with a much higher impact warning message such as pictorial health warnings that elicits strong emotional reactions which are significantly more effective [43].

It is evident that the risk communication is a very important step when designing the warning content. While in our study the warning message were randomly displayed to the user, it would be very interesting to see the impact of the warning message on the frequency of repeated uses by assigning a particular impact level warning (e.g. medium) to each participant, which would then be increased to reach a high level impact (e.g. the pictorial health-warning message).

Third, we hypothesized that the warning message will reduce the duration of both first and repeated software uses. Our study has an important finding where the warning message reduces the duration of both first and repeated software uses. This finding offers important empirical contribution supporting the restrictive deterrence model [18, 20]. The duration reduction can be further explained by the fact that users want to reduce their exposure for longer periods as they believe that some malicious actions could be committed by the software against them. In that context, users will simply try to avoid the punishment [44] and react to the sanction threats originating from the environment [24].

Overall, the results of our study suggest that the warning message affects the human behavior. Knowing that the human factor is usually the weakest link in the security chain [45] and considered to be the last line of defense against security risks [3], we argue that the importance of the warning message in that context becomes even higher.

However, in order to fully understand the warning message effectiveness other factors need to be taken into consideration. One of them relates to the fact that

the language in existing warnings is not as clear as it could be [16]. This is an interesting insight, as it could be that the warning communication is simply not what user expects to read. For example, even a simple word such as 'malware' may not be understandable by general user population.

Moreover, another aspect that could further explain our results relates to the organizational trust. Users will generally have more trust in software that was issued by organizations that have certain reputation. Clearly, the level of compliance will depend on the level of institutional trust [46]. All these factors, to certain extent, can influence the effectiveness of the sanction threat communicated by the warning message.

5.1. Theoretical contribution

Our study attempted to uncover the facets of the human-computer technology realm, and in particular the interaction between the human-warning message-computer system, which is particularly interesting as it is not only bringing more clarity to this relationship, but it also suggests future avenues to study the antecedents of human behavior. We empirically investigated the effects of a sanction threat on the occurrence, progression and duration of the software use. By doing so, not only did we support the restrictive deterrence model [18] but also brought important theoretical insights. Indeed, the deterrence effect is particularly visible and pronounced in the time continuum (specifically, after three seconds) which suggests that a restrictive deterrence is quite effective but is also reaching its limits, as it does not explain why the frequency of repeated uses is not reduced.

In addition, we advanced the fact the interaction between the human and the computer system, which is mediated by the computer warning, is a very complex model that requires further theorizing. Our study, offers important insights toward this direction. Thanks to the unique setting we used in this study and the fact that we did not actively recruit any participants, enabled us to investigate more accurate relationships in the human-computer system interaction.

5.2. Practitioner contribution

In this study we used software application which was displaying computer warnings to users aiming at understanding their behaviors in the decision making process. For the end user this 'displaying' process was transparent as it could not be easily identified if the warning message was coming from the software itself or from the operating system.

There are several ways our study could contribute to the existing challenges related to the information systems security. When it comes to installing new software, plugging-in hardware (e.g. external hard disk, USB key, etc.), or doing any action that can put in risk user's data or privacy integrity, we believe that the operating system (or any other automated computer technology related mechanism) should better interact with the user by displaying a more efficient warning message. And this does not relate only to the malware context, which we explore in this study, but it can be applied to all contexts where risks may be present. For example, we suggest that computer technology should start the interaction with the user when USB key is inserted into the computer and warn the user about the underlying risks.

5.3. Limitations and future research

Our study has several limitations. Due to the nature of the research design, as we did not recruit any of the participants, we also could not collect any demographics from the participants nor to follow up with additional surveys to better understand who the users are and get better understanding of their technical skills, which could have some influence on the results interpretation. Another limitation is that among the users that constantly ignored the warnings, we could not check if they knew that the warning was coming from the software and not from the operating system, which was our intention.

For future research, it would be interesting to extend the participant sample to business users (organizations) or students as these two populations could bring interesting new insights about the deterrent effects of warning messages.

Also, in our study we used a restrictive deterrence model as the theoretical foundation and built the warning message content using some of the psychological cues (e.g. risk consequences), but future studies could extend on this by using theories used from other disciplines which could be applied in the information systems context (e.g. health belief model or accountability theory).

6. Conclusion

Our study represents one of the first attempts at addressing an important issue: what are the effects of warnings on user's behavior in preventing malware use? The results suggest that the warning message affects the human behavior. Supported by the

restrictive deterrent concept and using psychological factors to build the warning messages, we found that in the presence of the warning message, the progression of the software use will be decreased and the duration of both first and repeated software uses will be reduced..

7. References

[1] Ponemon Institute, "The Cost of Malware Containment", in (Editor, 'ed.'^'eds.'): Book The Cost of Malware Containment, 2015

[2] Mcaffee, "Net Losses: Estimating the Global Cost of Cybercrime", in (Editor, 'ed.'^'eds.'): Book Net Losses: Estimating the Global Cost of Cybercrime, 2014

[3] Anderson, B.B., Vance, A., Kirwan, B., Eargle, D., and Howard, S., "Why Users Habituate to Security Warnings: Insights from Fmri", 2014 IFIP 8.11 Dewald Roode Security Workshop, 2014

[4] Wogalter, M.S., "Purposes and Scope of Warnings", Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ, 2006, pp. 3-9.

[5] Silic, M., Barlow, J., and Ormond, D., "Warning! A Comprehensive Model of the Effects of Digital Information Security Warning Messages", The 2015 Dewald Roode Workshop on Information Systems Security Research, IFIP, 2015, pp. 1-32.

[6] Kalsher, M.J., and Williams, K.J., "Behavioral Compliance: Theory, Methodology, and Results", Handbook of warnings, 2006, pp. 313-331.

[7] Egilman, D., and Bohme, S., "A Brief History of Warnings", Handbook of Warnings. Lawrence Erlbaum Associates, Mahwah, NJ, 2006, pp. 35-48.

[8] Anderson, B.B., Kirwan, C.B., Jenkins, J.L., Eargle, D., Howard, S., and Vance, A., "How Polymorphic Warnings Reduce Habituation in the Brain—Insights from an Fmri Study", Proc. of CHI'15, 2013,

[9] Bravo-Lillo, C., Cranor, L.F., Downs, J., Komanduri, S., and Sleeper, M., "Improving Computer Security Dialogs": Human-Computer Interaction–Interact 2011, Springer, 2011, pp. 18-35.

[10] Silic, M., "Understanding Colour Impact on Warning Messages: Evidence from Us and India", in (Editor, 'ed.'^'eds.'): Book Understanding Colour Impact on Warning Messages: Evidence from Us and India, ACM, 2016, pp. 2954-2960.

[11] Silic, M., Silic, D., and Oblakovic, G., "Restrictive Deterrence: Impact of Warning Banner Messages on Repeated Low-Trust Software Use", 18th International Conference on Enterprise Information Systems (ICEIS 2016), 2016, pp. 435-442.

[12] Silic, M., Silic, D., and Oblakovic, G., "The Effects of Colour on Users' Compliance with Warning Banner Messages across Cultures", ECIS 2016, 2016

[13] Silic, M., and Cyr, D., "Colour Arousal Effect on Users' Decision-Making Processes in the Warning Message Context", in (Editor, 'ed.'^'eds.'): Book Colour Arousal Effect on Users' Decision-Making Processes in the Warning Message Context, Springer, 2016, pp. 99-109.

[14] Akhawe, D., and Felt, A.P., "Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness", in (Editor, 'ed.'^'eds.'): Book Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness, USENIX Association, 2013, pp. 257-272.

[15] Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., and Cranor, L.F., "Crying Wolf: An Empirical Study of Ssl Warning Effectiveness", in (Editor, 'ed.'^'eds.'): Book Crying Wolf: An Empirical Study of Ssl Warning Effectiveness, 2009, pp. 399-432.

[16] Modic, D., and Anderson, R., "Reading This May Harm Your Computer: The Psychology of Malware Warnings", Computers in Human Behavior, 41(2014, pp. 71-79.

[17] Bravo-Lillo, C., Cranor, L.F., Downs, J., Komanduri, S., Reeder, R.W., Schechter, S., and Sleeper, M., "Your Attention Please: Designing Security-Decision Uis to Make Genuine Risks Harder to Ignore", in (Editor, 'ed.'^'eds.'): Book Your Attention Please: Designing Security-Decision Uis to Make Genuine Risks Harder to Ignore, 2013, pp. 1-18.

[18] Gibbs, J.P., Crime, Punishment, and Deterrence, Elsevier New York, 1975.

[19] Maimon, D., Alper, M., Sobesto, B., and Cukier, M., "Restrictive Deterrent Effects of a Warning Banner in an Attacked Computer System", Criminology, 52(1), 2014, pp. 33-59.

[20] Jacobs, B.A., "Deterrence and Deterrability*", Criminology, 48(2), 2010, pp. 417-441.

[21] Straub, D.W., and Welke, R.J., "Coping with Systems Risk: Security Planning Models for Management Decision Making", Mis Quarterly, 1998, pp. 441-469.

[22] Maimon, D., Alper, M., Sobesto, B., and Cukier, M., "Restrictive Deterrent Effects of a Warning Banner in an Attacked Computer System", Criminology, 52(2014, pp. 33-59.

[23] Jacobs, B.A., "Crack Dealers' Apprehension Avoidance Techniques: A Case of Restrictive Deterrence", Justice Quarterly, 13(3), 1996, pp. 359-381.

[24] Jacobs, B.A., and Cherbonneau, M., "Auto Theft and Restrictive Deterrence", Justice quarterly, 31(2), 2014, pp. 344-367.

[25] Gallupe, O., Bouchard, M., and Caulkins, J.P., "No Change Is a Good Change? Restrictive Deterrence in Illegal Drug Markets", Journal of Criminal Justice, 39(1), 2011, pp. 81-89.

[26] Beauregard, E., and Bouchard, M., "Cleaning up Your Act: Forensic Awareness as a Detection Avoidance Strategy", Journal of Criminal Justice, 38(6), 2010, pp. 1160-1166.

[27] Cranor, L.F., "A Framework for Reasoning About the Human in the Loop", UPSEC, 8(2008, pp. 1-15.

[28] Wogalter, M.S., "Communication-Human Information Processing (C-Hip) Model", Handbook of warnings, 2006, pp. 51-61.

[29] Bravo-Lillo, C., Cranor, L.F., Downs, J.S., and Komanduri, S., "Bridging the Gap in Computer Security Warnings: A Mental Model Approach", IEEE Security & Privacy, 9(2), 2011, pp. 0018-0026.

[30] Bauer, L., Bravo-Lillo, C., Cranor, L., and Fragkaki, E., "Warning Design Guidelines", in (Editor, 'ed.'^'eds.'): Book Warning Design Guidelines, Carnegie Mellon University, Pittsburgh, PA, 2013

[31] Egelman, S., and Schechter, S., "The Importance of Being Earnest [in Security Warnings]": Financial Cryptography and Data Security, Springer, 2013, pp. 52-59.

[32] Rsa Conference. (2014). Latest Guidelines for Malware Detection Retrieved April 2015, from http://www.rsaconference.com/blogs/latest-guidelines-for-malware-detection

[33] Apple, "How to Avoid or Remove Mac Defender Malware in Mac Os X V10.6 or Earlier", in (Editor, 'ed.'^'eds.'): Book How to Avoid or Remove Mac Defender Malware in Mac Os X V10.6 or Earlier, 2015

[34] Wogalter, M.S., and Laughery, K.R., "Warning! Sign and Label Effectiveness", Current Directions in Psychological Science, 1996, pp. 33-37.

[35] Silic, M., and Back, A., "Shadow It–a View from Behind the Curtain", Computers & Security, 45(2014, pp. 274-283.

[36] Kaplan, E.L., and Meier, P., "Nonparametric Estimation from Incomplete Observations", Journal of the American statistical association, 53(282), 1958, pp. 457-481.

[37] Rich, J.T., Neely, J.G., Paniello, R.C., Voelker, C.C., Nussenbaum, B., and Wang, E.W., "A Practical Guide to Understanding Kaplan-Meier Curves", Otolaryngology-Head and Neck Surgery, 143(3), 2010, pp. 331-336.

[38] Box-Steffensmeier, J.M., and Jones, B.S., Event History Modeling: A Guide for Social Scientists, Cambridge University Press, 2004.

[39] Egelman, S., Acquisti, A., Molnar, D., Herley, C., Christin, N., and Krishnamurthi, S., "Please Continue to Hold an Empirical Study on User Tolerance of Security Delays", 2010,

[40] Bravo-Lillo, C., Cranor, L.F., Downs, J.S., and Komanduri, S., "Bridging the Gap in Computer Security Warnings: A Mental Model Approach", IEEE Security & Privacy, 9(2), 2011, pp. 18-26.

[41] Sheng, S., Wardman, B., Warner, G., Cranor, L., Hong, J., and Zhang, C., "An Empirical Analysis of Phishing Blacklists", in (Editor, 'ed.'^'eds.'): Book An Empirical Analysis of Phishing Blacklists, California, USA, 2009

[42] Wu, M., Miller, R.C., and Garfinkel, S.L., "Do Security Toolbars Actually Prevent Phishing Attacks?", in (Editor, 'ed.'^'eds.'): Book Do Security Toolbars Actually Prevent Phishing Attacks?, ACM, 2006, pp. 601-610.

[43] Hammond, D., "Health Warning Messages on Tobacco Products: A Review", Tobacco control, 2011, pp. tc. 2010.037630.

[44] Stafford, M.C., and Warr, M., "A Reconceptualization of General and Specific Deterrence", Journal of Research in Crime and Delinquency, 30(2), 1993, pp. 123-135.

[45] Workman, M., "A Test of Interventions for Security Threats from Social Engineering", Information Management & Computer Security, 16(5), 2008, pp. 463-483.

[46] Wu, J.-J., and Tsang, A.S., "Factors Affecting Members' Trust Belief and Behaviour Intention in Virtual Communities", Behaviour & Information Technology, 27(2), 2008, pp. 115-125.